

Recogida, Tratamiento de datos y Distribución de Frecuencias

Tema 5

Esquema introductorio

- 1. Bibliografía, Autoevaluación y otros recursos
- 2. Recogida y Tratamiento de datos:
 - 2.1. Cuestionarios / Hojas de registro
 - 2.2. Matriz de datos
- 3. Distribución de Frecuencias:
 - 3.1. Qué es una distribución de frecuencias
 - 3.2. Tipos de frecuencias:
 - Simples
 - Acumuladas

1. Bibliografía, Autoevaluación y otros recursos

En este documento puedes encontrar un esquema con los contenidos básicos de este tema. Para prepararlo con más detenimiento puedes seguir el manual:

 Camarero, L. (Coord.): ***Estadística para la Investigación Social***. Madrid. Garceta. (Capítulo III)

Para hacer una autoevaluación sobre lo aprendido utiliza el documento:

 Autoeval_5

Para ver un documental sobre los resultados de un estudio en el que aparecen, de forma aplicada, algunos contenidos de este tema (principalmente frecuencias en porcentajes) :

 Pincha [aquí](#).

2. Recogida y Tratamiento de datos

- El análisis estadístico de los datos requiere de dos pasos previos:
 - RECOGIDA DE DATOS (en un **Cuestionario** o en una **Hoja de registro**)
 - ORDENACIÓN Y CLASIFICACIÓN de los datos (en una **Matriz de datos**)

2.1. Cuestionarios y Hojas de registro

- **Cuestionario:** es el soporte material a través del cual se aplica una encuesta. Y hace referencia al instrumento concreto de recopilación de datos, rigurosamente estandarizado a través de preguntas y respuestas.

(El término encuesta es más amplio que el de cuestionario y supone una técnica de investigación completa: diseño del cuestionario, aplicación del cuestionario -trabajo de campo-, tabulación y análisis de datos).

- **Hoja de registro:** es el soporte material (historial, ficha, expediente...) a través del cual se anota, de forma sistemática, información relevante sobre los usuarios o cualquier otra unidad de análisis.

2.1. Cuestionarios y Hojas de registro

- En un cuestionario o en una hoja de registro, y generalmente en forma de preguntas, se incluyen una serie de **variables** (características) que se desean conocer acerca de una población.
- Cada variable incluida en el cuestionario u hoja de registro, cuenta con una serie de **categorías o valores**.
- A su vez a cada categoría se le **codifica** con un número (código).

La asignación de números (códigos) a las categorías, como 1 a ser varón, 2 a ser mujer, es arbitraria. Son simplemente códigos que nos permiten posteriormente hacer cálculos matemáticos.

2.1. Cuestionarios y Hojas de registro (1er ejemplo)

- Por ejm, el cuestionario de la “**Encuesta de Condiciones de vida**” (INE, 2009) incluye la siguiente pregunta (variable), cuyas categorías aparecen especificadas y codificadas así:

P. 8 ¿Ha realizado alguna vez algún trabajo, ya sea asalariado, por su cuenta o como ayuda familiar?

| | | |
|---------|---|-----------|
| Sí..... | 1 | |
| No..... | 6 | ← Códigos |

↑
Categorías

2.1. Cuestionarios y Hojas de registro (1er ejemplo)

 En la “Encuesta de Condiciones de Vida” (INE, 2009) se incluyen, entre otras, variables acerca de:

“Situación económica”, “pobreza”, “privación”, “protección mínima”, “igualdad de trato”, “empleo”, “actividad”, “formación profesional permanente”, “migraciones laborales”, “jubilaciones”, “pensiones”, “status socioeconómico de las personas de edad”, “nivel de formación” y “efectos sobre la condición socioeconómica”.

- El **Objetivo** de la ***Encuesta de Condiciones de Vida*** es estudiar el nivel socio-económico, las condiciones del mercado de trabajo y la cohesión social en relación con los requerimientos de información de las políticas activas de la UE en estos ámbitos y con sus efectos para la población.
- En la actualidad esta encuesta es un instrumento clave para el estudio de la pobreza y la exclusión social.

2.1. Cuestionarios y Hojas de registro (1er ejemplo)

ECV_2009.pdf - Adobe Reader

Archivo Edición Ver Documento Herramientas Ventana Ayuda

1 / 21 143% Buscar

INE

INSTITUTO NACIONAL DE ESTADÍSTICA

Encuesta de Condiciones de Vida. ECV(6). 2009

Cuestionario Individual

A. Identificación

1. Identificación del hogar

Provincia _____ | | |

Nº de orden de la sección _____ | | | |

Hogar/Vivienda _____ | | | | | | |


Rama/Nº de orden del hogar _____ | |

3

ENTREVISTADOR: Antes de comenzar a cumplimentar este cuestionario, por favor, lea al informante la siguiente introducción: "A continuación voy a hacerle unas preguntas referentes a usted sobre trabajo, ingresos, salud, estudios, etc."

 Cuestionario de la *Encuesta de Condiciones de Vida* (INE, 2009).

2.1. Cuestionarios y Hojas de registro (2º ejemplo)

- En la actualidad cualquier institución o entidad dedicada a la Intervención Social cuenta con **Hojas de registro informatizadas** en las que se anotan las características de sus usuarios.
- El programa informático **SIUSS (Sistema de Información de Usuarios de Servicios Sociales)** del Ministerio de Sanidad, Política Social e Igualdad es un soporte documental que permite la recogida de datos básicos de los usuarios de Servicios Sociales. Este programa funciona a modo de “Gran hoja de registro” en la que diferentes comunidades autónomas vuelcan sus datos.  **Para más información sobre los objetivos y la descripción del SIUSS, así como los principales resultados del análisis estadístico de 2008. pincha [aquí](#).**
- El Ayuntamiento de Madrid dispone del programa **SELENE**. En él los profesionales de las Unidades de Trabajo Social anotan datos sobre las características de los usuarios, situación socio-familiar, intervención social, valoración, etc.

2.2 MATRIZ DE DATOS

- La grabación, ordenación y clasificación de los datos recogidos en un cuestionario / hoja de registro da lugar a la matriz de datos.
- Una matriz de datos tiene "n" filas y "m" columnas ("n" x "m").
 - En filas se colocan los casos o unidades (sujetos entrevistados).
 - En columnas se colocan las variables.

2.2. MATRIZ DE DATOS

| n filas | n columnas | | | |
|---------|------------|------------|------------|------------|
| | | Variable_1 | Variable_2 | Variable_n |
| | Caso_1 | | | |
| | Caso_2 | | | |
| | Caso_n | | | |

2.2. Matriz de datos (1er ejemplo)

- La siguiente matriz de datos contiene 4 casos de los que se ha obtenido información en relación a dos **variables** ("Sexo" y "Estado civil").
- Previamente se han establecido los **categorías/valores...** de las variables y su respectivos **códigos**:
 - En la variable "sexo":
 - "Varón" (1)
 - "Mujer" (2)
 - La variable Estado civil:
 - "Soltero" (1)
 - "Casado" (2)
 - "Pareja de hecho" (3)
 - "Separado /divorciado" (4)
 - "Viudo" (5)

2.2. Matriz de datos (1er ejemplo)

| | Sexo | Estado civil |
|---|-----------|---------------------|
| 1 | 1 (Varón) | 2 (Casado) |
| 2 | 2 (Mujer) | 3 (Pareja de hecho) |
| 3 | 2 (Mujer) | 1 (Soltero) |
| 4 | 2 (Mujer) | 1 (Soltero) |

En esta matriz de datos podemos ver que:

- ✓ El caso 1 es un “Varón”, “Casado”.
- ✓ El caso 2 es una “Mujer”, viviendo en “Pareja de Hecho”.
- ✓ El caso 3 es una “Mujer”, “Soltera”.
- ✓ El caso 4 es una “Mujer”, “Soltera”.

2.2. Matriz de datos (2º ejemplo)

RED.sav [Conjunto_de_datos1] - Editor de datos SPSS

Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana ?

21 : Visible: 10 de 10 variables


| | a1_1 | a1_2 | b5 | b3 | b12 | j92 | f55 | i89 | i78 | k105 | var | var | var | var | var | var |
|----|------|------|----|----|-----|-----|-----|-----|-----|------|-----|-----|-----|-----|-----|-----|
| 1 | 20 | 69 | 1 | 1 | | 1 | 7 | 2 | 1 | 4 | 3 | | | | | |
| 2 | 20 | 69 | 6 | 1 | | 2 | 1 | 2 | 6 | 2 | 1 | | | | | |
| 3 | 20 | 69 | 6 | 1 | | 2 | 5 | 2 | 1 | 3 | 3 | | | | | |
| 4 | 20 | 69 | 1 | 1 | | 1 | 1 | 2 | 1 | 3 | 3 | | | | | |
| 5 | 20 | 69 | 1 | 6 | | 1 | 3 | 2 | 1 | 3 | 3 | | | | | |
| 6 | 48 | 20 | 6 | 1 | | 2 | 1 | 2 | 1 | 2 | 3 | | | | | |
| 7 | 48 | 20 | 6 | 1 | | 2 | 2 | 2 | 6 | 4 | 1 | | | | | |
| 8 | 48 | 20 | 6 | 1 | | 2 | 1 | 2 | 6 | 1 | 4 | | | | | |
| 9 | 48 | 20 | 1 | 1 | | 1 | 7 | 4 | 6 | 4 | 3 | | | | | |
| 10 | 48 | 20 | 1 | 1 | | 1 | 1 | 3 | 6 | 4 | 2 | | | | | |
| 11 | 48 | 20 | 1 | 1 | | 1 | 4 | 3 | 6 | 3 | 2 | | | | | |
| 12 | 48 | 20 | 6 | 1 | | 2 | 1 | 1 | 6 | 2 | 3 | | | | | |
| 13 | 20 | 69 | 1 | 1 | | 1 | 1 | 4 | 1 | 3 | 3 | | | | | |
| 14 | 48 | 20 | 6 | 1 | | 2 | 1 | 1 | 1 | 2 | 3 | | | | | |
| 15 | 1 | 59 | 1 | 6 | | 1 | 1 | 2 | 1 | 2 | 1 | | | | | |
| 16 | 1 | 59 | 6 | 1 | | 2 | 5 | 1 | 1 | 2 | 1 | | | | | |
| 17 | 1 | 59 | 6 | 1 | | 2 | 1 | 1 | 1 | 2 | 1 | | | | | |
| 18 | 20 | 69 | 1 | 1 | | 1 | 1 | 2 | 1 | 2 | 1 | | | | | |
| 19 | 20 | 69 | 1 | 1 | | 1 | 6 | 2 | 1 | 3 | 3 | | | | | |
| 20 | 20 | 69 | 6 | 6 | | 2 | 1 | 2 | 1 | 4 | 2 | | | | | |
| 21 | 20 | 69 | 6 | 1 | | 2 | 3 | 1 | 6 | 2 | 1 | | | | | |
| 22 | 20 | 69 | 1 | 6 | | 1 | 6 | 2 | 6 | 5 | 3 | | | | | |
| 23 | 1 | 59 | 1 | 1 | | 1 | 1 | 4 | 6 | 3 | 2 | | | | | |
| 24 | 48 | 20 | 1 | 1 | | 1 | 2 | 2 | 1 | 4 | 3 | | | | | |
| 25 | 20 | 45 | 1 | 6 | | 1 | 7 | 2 | 6 | 3 | 3 | | | | | |
| 26 | 20 | 45 | 1 | 6 | | 1 | 4 | 2 | 6 | 2 | 3 | | | | | |
| 27 | 48 | 20 | 1 | 1 | | 1 | 1 | 4 | 6 | 2 | 2 | | | | | |
| 28 | 48 | 20 | 1 | 1 | | 1 | 1 | 1 | 6 | 2 | 2 | | | | | |
| 29 | 48 | 20 | 1 | 1 | | 1 | 1 | 1 | 6 | 2 | 3 | | | | | |
| 30 | 48 | 20 | 1 | 1 | | 1 | 5 | 2 | 6 | 3 | 3 | | | | | |
| 31 | 48 | 20 | 6 | 1 | | 2 | 3 | 2 | 6 | 3 | 3 | | | | | |
| 32 | 20 | 69 | 1 | 6 | | 1 | 1 | 2 | 1 | 3 | 1 | | | | | |

Vista de datos Vista de variables /

SPSS El procesador está preparado

Inicio DATOS_ASSIGNAT... RED.sav [Conjunt... Documento1 - Mic...

ES 20:27

 Puedes encontrar un extracto de una matriz de datos de un estudio sobre **“Usuarios de la red de Atención Social Especializada de Servicios Sociales”** (Rubio, M.J. , 2009) en el archivo adjunto RED.pdf

2.2. Matriz de Datos: Paquetes estadísticos para la confección de una matriz de datos

- Algunos de los principales paquetes estadísticos tanto para la confección de una **matriz de datos** como para su posterior **análisis estadístico** son:
 - SPSS
 - SAS
 - STATA
 - SPHINX SURVEY
 - Gandía Barwin
 - También EXCEL permite construir matrices de datos, así como realizar sencillos análisis estadísticos.

3. Distribución de Frecuencias

- Una vez realizada la matriz de datos el proceso a seguir consiste en:
 - Agrupación de los datos que se repiten (Frecuencias).
 - Presentación de la tabla estadística con sus distintas frecuencias.
 - Lectura/análisis de la tabla de frecuencias.

3.1. Qué es una distribución de frecuencias

- Se denomina **FRECUENCIA** al número de casos que contiene cada categoría / valor de una variable.
- Se denomina **DISTRIBUCIÓN DE FRECUENCIAS** a la forma que adopta un suceso que se repite y que previamente se ha organizado en categorías o clases.

3.1 Distribución de Frecuencias

(1er. Ejemplo)

En una ficha de seguimiento (Hoja de registro) acerca de 20 familias, se ha registrado el número de hijos (variable: “Número de hijos”):

4, 3, 4, 2, 0, 4, 1, 3, 2, 3, 4, 3, 2, 3, 1, 2, 4, 3, 2, 3,

Una vez ordenados los datos, tenemos:

0, 1, 1, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4

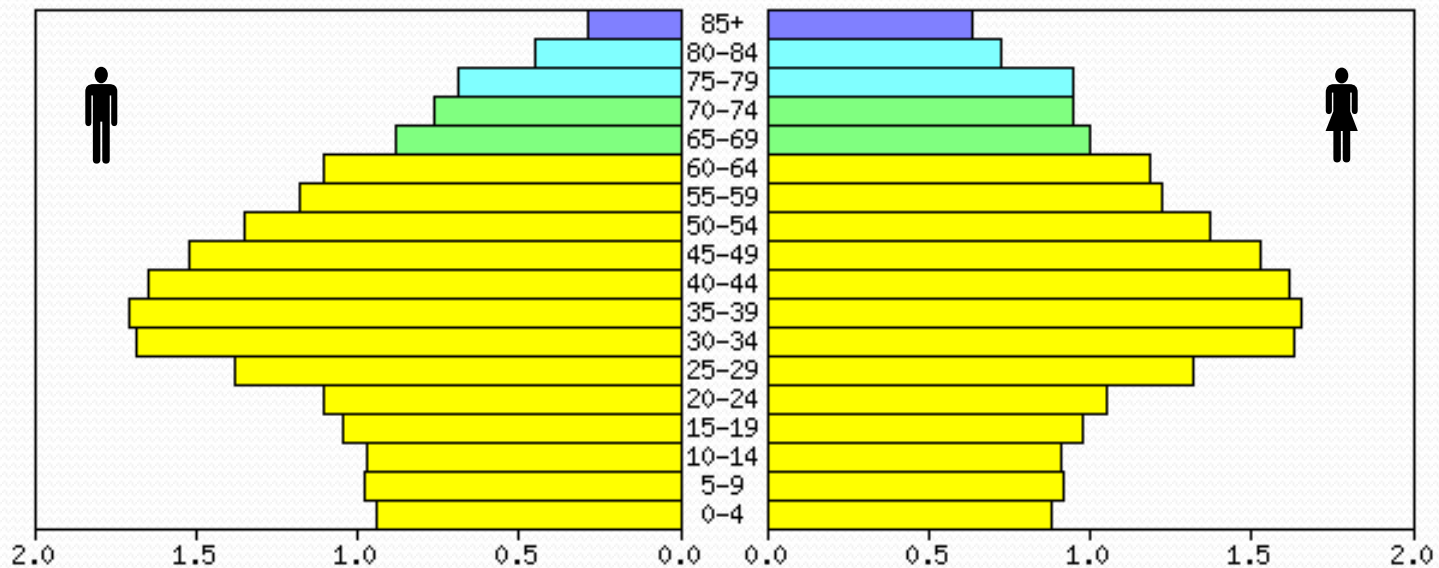
Y ahora agrupamos los datos en **frecuencias** y obtenemos la **distribución de frecuencias** de la variable “Número de hijos”:

| Nº de hijos | Frecuencias |
|-------------|-------------|
| 0 | 1 |
| 1 | 2 |
| 2 | 5 |
| 3 | 7 |
| 4 | 5 |
| | 20 |

Distribución de Frecuencias. 2º Ejemplo

(INE, 2009)

- Las pirámides de población son representaciones gráficas de las **distribuciones de frecuencias** de dos **variables**: Sexo y Edad.



3.2. Tipos de Frecuencias

- Frecuencias Simples:
 - Absolutas
 - Relativas
- Frecuencias Acumuladas:
 - Absolutas
 - Relativas

3.2. Tipos de Frecuencias

Frecuencias Simples:

- **Frecuencia absoluta (n_i):** es el número de veces que se repite un suceso en cada categoría/valor de la variable.
 - El número total de casos (N) es igual a la suma de las frecuencias absolutas: $N = n_1 + n_2 + \dots + n_n$
- **Frecuencia relativa (f_r):** es el cociente entre la frecuencia absoluta de cada categoría y el total de los casos. (Sirve para hacer una comparación numérica entre la parte y el todo. Entre el número de casos de cada categoría y el total de los casos).

Las frecuencias **relativas** pueden expresarse de dos formas:

- **En proporciones.** La suma de todas las frecuencias relativas es igual a 1.
- **En porcentajes** ($f_r \times 100$). La suma de todas las frecuencias relativas es igual a 100.

3.2. Tipos de Frecuencias. Frecuencias simples (1er ejemplo)

TABLA 1

| Nº de hijos | n_i | f_r | % |
|--------------------|-------------------------|-------------------------|----------|
| 0 | 1 | 0,05 | 5 |
| 1 | 2 | 0,10 | 10 |
| 2 | 5 | 0,25 | 25 |
| 3 | 7 | 0,35 | 35 |
| 4 | 5 | 0,25 | 25 |
| | N=20 | 1 | 100 |

Ejemplo: ficha de seguimiento (Hoja de registro) acerca de 20 familias, en la que se ha anotado el número de hijos (variable: “Número de hijos”).

3.2. Tipos de Frecuencias. Frecuencias simples (1er ejemplo). Lectura de Tablas

- Una vez clasificada la variable “Nº de hijos” en la tabla anterior, se puede ver el reparto de las familias para cada valor. Por ejemplo, 5 familias tiene 2 hijos, 3 familias tienen 7 hijos, sólo 1 familia no tiene hijos...
- Si queremos comparar los valores de la variable, y saber qué parte del total de las familias contiene un determinado valor, por ejemplo, cuánto representan las familias con 3 hijos respecto del total, nos fijamos en las proporciones o en los porcentajes: el 35% de las familias estudiadas tiene 3 hijos.

$$0,35 = 0,35 \times 100 = 35\%$$

3.2. Tipos de Frecuencias

Frecuencias acumuladas:

- **Frecuencia absoluta acumulada (N_a)**: la frecuencia absoluta de cada categoría/valor es la suma de la frecuencia de cada categoría con las anteriores. (La frecuencia absoluta acumulada de la primera categoría/valor es la frecuencia de esa misma categoría).
- **Frecuencia relativa acumulada (F_{ra})**: es la suma de la frecuencia relativa de cada categoría con las anteriores. La frecuencia relativa acumulada de la primera categoría/valor es la frecuencia relativa de esa misma categoría).
- Las Frecuencias relativas acumuladas (F_{ra}) también pueden expresarse en forma de **porcentajes** ($\% F_{ra}$).

3. Tipos de Frecuencias. Frecuencias Acumuladas (1er ejemplo)

TABLA 2

| Nº de hijos | n_i | N_a | (F_{ra}) | % |
|-------------|-------|-------|------------|-----|
| 0 | 1 | 1 | 0,05 | 5 |
| 1 | 2 | 3 | 0,15 | 15 |
| 2 | 5 | 8 | 0,40 | 40 |
| 3 | 7 | 15 | 0,75 | 75 |
| 4 | 5 | 20 | 1 | 100 |
| | N=20 | | | |

Ejemplo: ficha de seguimiento (Hoja de registro) acerca de 20 familias, en la que se ha anotado el número de hijos (variable: “Número de hijos”):

3. Tipos de Frecuencias. Frecuencias Acumuladas (1er ejemplo). Lectura de Tablas

- En la Tabla de frecuencias absolutas anterior podemos ver que sólo 3 familias (15%) tienen 1 o ningún hijo.
- Pero también podemos calcular las frecuencias acumuladas de forma inversa, de mayor a menor, obteniendo la Tabla que aparece en la página siguiente. Y en la que de forma directa podemos observar que más de la mitad de las familias estudiadas (60%) tienen 3 o más hijos, y el 85% de las familias tiene 2 o más hijos.

$$0,25+0,35=0,60 \times 100=60\%$$

$$0,25+0,35+0,25=0,85 \times 100=85\%$$

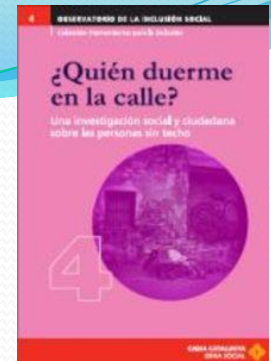
3. Tipos de Frecuencias. Frecuencias Acumuladas (1er ejemplo)

TABLA 3

| Nº de hijos | n_i | N_a | (F_{ra}) | % |
|-------------|-------|-------|------------|-----|
| 0 | 1 | 20 | 1 | 100 |
| 1 | 2 | 19 | 0,95 | 95 |
| 2 | 5 | 17 | 0,85 | 85 |
| 3 | 7 | 12 | 0,60 | 60 |
| 4 | 5 | 5 | 0,25 | 25 |
| | N=20 | | | |

Ejemplo: ficha de seguimiento (Hoja de registro) acerca de 20 familias, en la que se ha anotado el número de hijos (variable: “Número de hijos”).

3.2. Tipos de Frecuencias (2º ejemplo)




- En siguiente ejemplo se muestran dos Tablas de frecuencias de un recuento nocturno de las Personas sin Techo que duermen en las calles de Barcelona, realizado en 2008.
- La investigación no sólo permitió dimensionar y caracterizar el problema, sino también poner rostro a estas personas , saber cómo son y concretar cuáles son sus necesidades.
- Contó con la colaboración de 708 voluntarios, que además de hacer operativo el recuento, entrevistaron a los Sin Techo. Lo que les permitió descubrir una realidad más allá de los lugares comunes y las imágenes estereotipadas, y contribuir a la sensibilización ciudadana sobre el problema.
- Puedes ver un video-resumen sobre él pinchando [aquí](#) . (El vídeo comienza apuntando que en las calles de Barcelona duermen 1878. Esta cifra es el resultado de sumar las personas que duermen literalmente en la calle con las que lo hacen en algún tipo de albergue o centro de emergencia destinado a tal fin).

3.2 Tipos de Frecuencias (2º ejemplo)

(Frecuencias simples)

TABLA 4

Esta tabla refleja la distribución de las personas que duermen al raso en las calles de Barcelona según los Distritos. En la primera columna, en frecuencias absolutas (n_i). En la segunda, en frecuencias relativas (f_r).

 Fuente: Cabrera, Rubio y Blasco (2008): *¿Quién duerme en la calle? Un Estudio social y ciudadano sobre Personas sin Hogar*. Barcelona. Caixa Catalunya. Puedes bajar este estudio pinchando [aquí](#).

| Distrito | Número de personas (n_i) | (f_r) |
|---------------------|---------------------------------|-----------|
| Eixample | 168 | 0,284 |
| Ciutat Vella | 151 | 0,255 |
| Sants-Montjuïc | 80 | 0,135 |
| Sant Martí | 68 | 0,115 |
| Sarrià-Sant Gervasi | 50 | 0,084 |
| Les Corts | 24 | 0,040 |
| Gràcia | 22 | 0,037 |
| Sant Andreu | 13 | 0,021 |
| Horta Guinardó | 9 | 0,015 |
| Nou Barris | 6 | 0,010 |
| Total | 591 | 1 |

3.2. Tipos de Frecuencias (2º ejemplo)

(Frecuencias simples y acumuladas)

TABLA 5

| Distrito | Número de personas (n_i) | % (f_r) | Acumulado (N_a) | % acumulado (F_{ra}) |
|---------------------|---------------------------------|----------------|------------------------|-----------------------------|
| Eixample | 168 | 28,4 | 168 | 28,4 |
| Ciutat Vella | 151 | 25,5 | 319 | 54,0 |
| Sants-Montjuïc | 80 | 13,5 | 399 | 67,5 |
| Sant Martí | 68 | 11,5 | 467 | 79,0 |
| Sarrià-Sant Gervasi | 50 | 8,5 | 517 | 87,5 |
| Les Corts | 24 | 4,1 | 541 | 91,5 |
| Gràcia | 22 | 3,7 | 563 | 95,3 |
| Sant Andreu | 13 | 2,2 | 576 | 97,5 |
| Horta Guinardó | 9 | 1,5 | 585 | 99,0 |
| Nou Barris | 6 | 1,0 | 591 | 100,0 |
| Total | 591 | 100 | | |

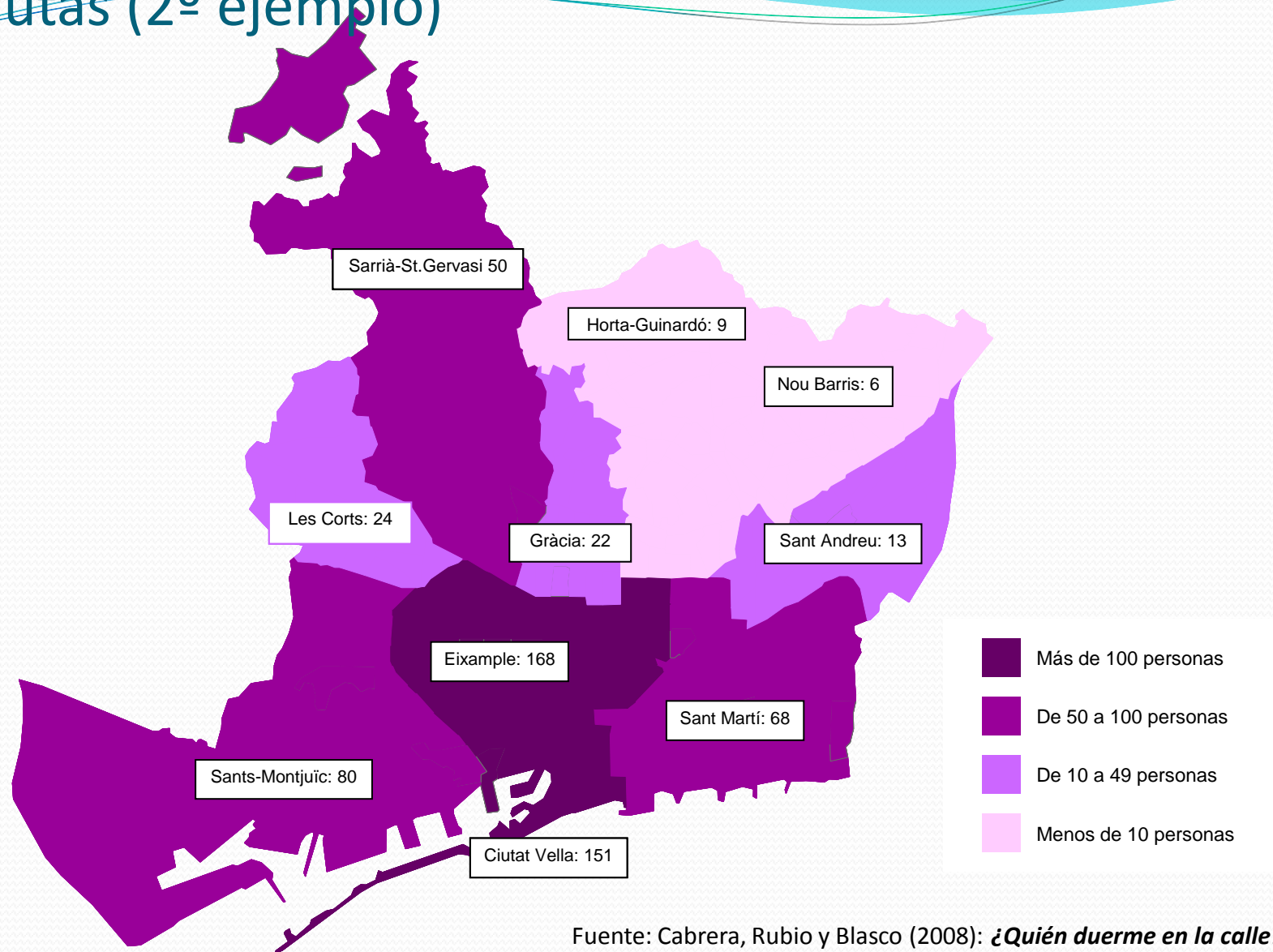
Fuente: Cabrera, Rubio y Blasco (2008): *¿Quién duerme en la calle? Un Estudio social y ciudadano sobre Personas sin Hogar*. Barcelona. Caixa Catalunya.

3.2. Tipos de Frecuencias (2º ejemplo)

Lectura de Tablas.

- Si observamos las dos tablas anteriores, podemos observar que, por ejemplo:
 - L'Eixample (28,4%) y Ciutat Vella (25,5%), son los dos distritos de Barcelona con más personas durmiendo en sus calles.
 - Prácticamente el 80% (79%) de las personas sin techo se concentran en únicamente cuatro distritos: Ciutat Vella, L'Eixample, Sants-Montjuïc y Sant Martí.

Representación gráfica a partir de Frecuencias simples absolutas (2º ejemplo)



Fuente: Cabrera, Rubio y Blasco (2008): *¿Quién duerme en la calle? Un Estudio social y ciudadano sobre Personas sin Hogar*. Barcelona. Caixa Catalunya.